

Enhanced Denoising System

Field of the Invention

5 The present invention relates to signal processing, and more particularly, to the correction of errors introduced into a signal by the transmission or processing of that signal.

Background of the Invention

10 The present invention can be more easily understood in terms of a simple exemplary system. Consider a telephone conversation in which a person talks into a microphone whose output is digitized and then transmitted to a second person via various telephone lines and switch systems. The speaker at the second person's location receives a sequence of digital values that are then played back to the second person. In general, the received sequence will
15 differ from the transmitted sequence because of errors introduced by the transmission system, digital-to-analog converters, and analog to digital converters. For example, noise in the transmission system results in some of the digital values in the transmitted sequence being altered. One goal of a denoising system is to remove as many of these noise errors as possible.

20

 The simple example discussed above is an example of a more general problem that is encountered in a wide range of applications. In general, an input digital signal that consists of a sequence of "symbols" is transmitted through a "communication link" and is received as an output digital signal at the output of the communication link. The output digital signal also
25 consists of a sequence of "symbols". Each of the symbols is chosen from a predetermined set of symbols, referred to as an alphabet. The output signal is assumed to be written in the same alphabet as the input signal.

 In the simplest case, the signals are binary signals in which the alphabet consists of the
30 symbols "0" and "1". In this case the input and output signals consist of a sequence of 0s and 1s. However, other alphabets are commonly used. For example, a digitized signal in which each symbol is represented by an integer between 0 and M-1 is commonly used in broadband

data transmission systems for connecting users to the Internet via a digital subscriber loop (DSL).

While the above examples refer to communication systems, it should be noted that this type of noise problem is present in a number of data processing systems. For example, the storage of data files on a magnetic disk drive can be viewed as the transmission of a digital signal through a communication link, the disk drive. The input signal is a sequence of symbols, e.g., bytes of data, which are chosen from a predetermined alphabet. In the case of byte data, each symbol has an integer value chosen from the set [0,1, ... ,255]. The retrieved file from the disk drive also consists of a sequence of symbols chosen from this set. The input signal symbols are processed by the electronics of the disk drive and stored in the form of localized magnetic fields that are read to generate the output signal. Noise in the digital to analog circuitry that converts the symbols to and from the magnetic fields introduces errors into the output signal. In addition, the magnetic fields can be altered during storage by random events that introduce additional errors.

Similarly, digital photography may be viewed as involving the transmission of a signal through a channel that corrupts the signal. In this case, the signal is the image, which is corrupted by noise in the photodetectors.

Summary of the Invention

The present invention includes a method and apparatus for processing a received digital signal that includes a sequence of symbols that has been corrupted by a channel to generate a processed digital signal. The method includes storing the received digital signal and receiving a partially corrected sequence of symbols that includes an output of a preliminary denoising system operating on the received digital signal. Information specifying a signal degradation function that measures the signal degradation that occurs if a symbol having the value I is replaced by symbol having a value J is utilized to generate a processed digital signal by replacing each symbol having a value I in a context of that symbol in the received digital signal with a symbol having a value J if replacement reduces a measure of overall signal degradation in the processed digital signal relative to the input digital signal as determined using the degradation function and the partially corrected sequence of symbols.

The method can be practiced on a dedicated apparatus or on a general purpose data processing system.

Brief Description of the Drawings

Figure 1 is a block diagram of a denoising system according to one embodiment of the present invention.

Figure 2 is a flow chart of the process used to determine the symbol value.

Figure 3 is a flow chart of the signal processing algorithm used in the second pass.

Detailed Description of the Preferred Embodiments of the Invention

The present invention provides a method for reducing the signal degradation resulting from the noise that is introduced into a digital signal when the signal is processed by a system that introduces noise errors. The processing system that introduces the noise will be referred to as the “channel” in the following discussion because such a system is analogous to a transmission channel over which the signal is sent.

Refer now to Figure 1, which is a block diagram of a denoising system 100 according to one embodiment of the present invention, operating on a signal 23 that has been corrupted by a channel 20. The channel operates on an input signal 21 comprising a sequence of symbols, y_1, y_2, \dots, y_n from a known alphabet to generate an output signal 23 that also comprises a sequence of symbols from that alphabet. The noisy output signal will be denoted by the sequence z_1, z_2, \dots, z_n . The noisy output signal symbols are also assumed to be from the same alphabet as the input signal symbols. That is, each symbol can take on a value from 0 to $M-1$, where M is an integer >1 . To simplify the following discussion, sequences of symbols will be denoted in boldface. For example, the sequence y_1, y_2, \dots, y_n will be denoted by \mathbf{y} .

It is assumed that a preliminary denoising system 120 operates on \mathbf{z} to generate a first approximation to a denoised signal 24, $\mathbf{r}=\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n$ by changing various members of the \mathbf{z} sequence in a manner that is not known to those receiving \mathbf{r} . Consider a subsequence of $2k+1$ symbols in \mathbf{z} that is centered about z_q . Here, k is an integer. The manner in which k is chosen
 5 will be discussed in more detail below. Denote this subsequence by $\mathbf{z}(q)$. That is $\mathbf{z}(q)=z_{q-k}, z_{q-k+1}, \dots, z_q, z_{q+1}, \dots, z_{q+k}$. The subsequence $\mathbf{z}(q)$ shall sometimes be referred to in what follows as the reference subsequence for index q . Assume that k is chosen such that this subsequence appears at a number of locations in \mathbf{z} . That is, $\mathbf{z}(p)=\mathbf{z}(q)$ for a number of different values of p . The present invention is based on the assumption that if the preliminary
 10 denoising system changes the value of z_q , it should also change the value of z_p in the same manner for each of the other occurrences of this subsequence.

The present invention examines the output of the preliminary denoising system and determines a value to be assigned to z_q and each of the z_p 's based on a measure of the signal
 15 degradation that occurs when a symbol is mistakenly replaced by another symbol. This resulting new sequence 22, \mathbf{z}' , is then output from the present invention. The present invention assumes that there is a quantified measure of the degradation introduced into the output signal by replacing a symbol having the value A in the input signal by a symbol having the value B in the output signal. The degradation may be different for different values of A
 20 and B . In the following discussion this degradation measure will be referred to simply as the "degradation" and denoted by $C(A,B)$.

In systems that utilize an alphabet that contains more than two symbols, $C(A,B)$ will often depend on the difference between A and B . For example, consider a digital signal that
 25 is generated by converting an analog time varying signal to a sequence of digital values utilizing an 8-bit analog-to-digital converter. The resulting digital signal is a sequence of symbols chosen from an alphabet having 256 symbols corresponding to the digital values 0 through 255. Assume that the output signal is to be converted back into an analog signal and played back to a human observer. The error in the output signal resulting from a symbol
 30 being altered by 1 is usually much less than the error resulting from a symbol being altered by a 2, and so on. Hence, the degradation function will depend on the amount by which the symbol is changed in this case.

The manner in which the present invention defines the correct symbol to use in place of z_q can be more easily understood with reference to Figure 2, which is a flow chart of the process used to determine the symbol value. The algorithm can be broken into two parts. In the first part, each subsequence centered at p for which $z(p)=z(q)$ is identified and the number of times the preliminary denoising system outputs each possible value for r_p is determined for each such p . Denote the number of times that r_p was assigned the value j by the preliminary denoising system by $N(j)$ for all of these values of p . The algorithm that implements the first part starts by initializing a number of variables as shown at 51. The algorithm then searches for each sequence for which $z(p)=z(q)$. For the current value of p , the algorithm tests $z(p)$ as shown at 52. If $z(p)=z(q)$, $N(r_p)$ is incremented as shown at 53. In either case, p is incremented to the next value as shown at 54, and the new value of p is tested to be sure that it is within the permissible range as shown at 55. If there are more subsequences to test, the process is repeated. When all of the subsequences have been examined, the algorithm proceeds to the second part.

In the second part of the algorithm, the counts from the first part are used to estimate the degradation that would result in the signal for the various possible choices of symbol values to which z_q could be changed. Consider the case in which z_q is changed to the value K . The algorithm computes the degradation estimate $D(K)$ as follows:

$$D(K) = \sum_{j=0}^{M-1} N(j)C(j,K) \quad (1)$$

as shown at 56. The algorithm then sets z'_q equal to K_{\min} , defined as the value of K for which $D(K)$ has the minimum value.

The manner in which the algorithm alters the output of the preliminary denoiser can be more easily understood with reference to a simple example. Consider the case in which the cost of making an error is the same for all errors, i.e., $C(I,J)=C_0$ for all I that are different from J . It should be noted that $C(I,I)=0$ for all I . In this case, $D(K)$ will be $S(K)C_0$, where $S(K)$ is the sum of $N(J)$ for J different from K . Now assume that $N(1) \gg N(J)$ for J different from 1. That is, in the vast majority of the cases, the preliminary denoiser substituted the value 1 for the symbol at the middle of each subsequence equal to $z(q)$ in the noisy signal. In this case, $D(K)$ will have its minimum value for $K=1$, since all of the other values of $D(K)$

will include $N(1)$ in the $S(K)$ term. Hence, for this degradation function, the algorithm of the present invention sets the output z'_p for all p for which $z(p) = z(q)$ to that value taken on by the majority of the r_p , the output of the preliminary denoiser, for such indices p .

The above-described embodiments utilized a $2k$ long sequence surrounding the symbol being processed to define the $2k+1$ symbol reference subsequence whose instances in \mathbf{z} and the corresponding symbols in \mathbf{r} are examined to determine the output symbol that is to be used in place of the symbol being processed. To simplify the following discussion of the more general cases, it is useful to define a "context" for the symbol being processed. Consider a symbol in the output signal. A subsequence of symbols having fixed values and in a predetermined location with respect to that symbol will be referred to as the "context" of that symbol. In the preceding example, the context of the symbol z_q was the k symbols on each side of z_q . Denote the k symbols on the left of z_q by $\mathbf{a} = a_1, a_2, \dots, a_k$ and the k symbols on the right of z_q by $\mathbf{b} = b_1, b_2, \dots, b_k$. Then the reference subsequence used to determine the replacement symbol for z_q can be written as $\mathbf{z}(q) = \mathbf{a}z_q\mathbf{b}$. It should be noted, however, that other contexts can be utilized in the present invention. For example, the sequence ending with the symbol z_q , i.e., $\mathbf{a}z_q$, could have been utilized. Similarly, the sequence beginning with z_q , i.e., $z_q\mathbf{b}$, could have been utilized. Furthermore, the lengths of the sequences \mathbf{a} and \mathbf{b} could be different.

In addition, contexts in which the sequences \mathbf{a} and/or \mathbf{b} have "wild cards" can also be utilized. That is, \mathbf{a} may be written in the form $a_1, a_2, \dots, \mathbf{a}_w, \dots, a_k$, where \mathbf{a}_w can be a string of symbols in which the symbols in the string can take on any value. Similarly, the symbols of the context do not need to be adjacent to the symbol being processed as long as they are in a predetermined location relative to that symbol. The above general definition of the context of a symbol and the induced reference subsequence applies also to multi-dimensional signals such as two-dimensional image data.

Refer again to Figure 1. In one embodiment of the present invention, \mathbf{z} and \mathbf{r} are read by denoising system 100. The \mathbf{z} sequence is stored in a memory 113 as it is received. For the purposes of this example, it will be assumed that the context of each symbol is the k symbols to the left of that symbol and the k' symbols to the right of that symbol. In the first pass, controller 111 stores the received sequence \mathbf{z} in a memory 113 as the symbols are received.

Controller 111 also makes a list of all subsequences of length $L = k + k' + 1$ in z . As each symbol is received, controller 111 examines the most recently received L symbols in z to determine the reference subsequence that has just been completed. Assume that the j^{th} symbol in the z sequence has just been received. This symbol completes $z(j - k' - 1)$, the reference subsequence associated with the symbol to be processed at $j - k' - 1$.

Controller 111 examines the sequences stored in memory 114 to determine if $z(j - k' - 1)$ has been received earlier. If not, controller 111 makes a new entry in memory 114 for the subsequence. The entry includes the L symbols that make up the subsequence and M counters for keeping track of the results from preliminary denoising system 120 for this sequence. Controller 111 then records the preliminary denoising system result in the appropriate counter. That is, controller 111 increments the counter corresponding to the symbol value $r_{j - k' - 1}$. When all of the symbols from both of the sequences z and r have been received and processed, the first pass is complete.

In the second pass, controller 111 sequentially goes through the stored z sequence and replaces each symbol with the symbol determined by the algorithm discussed above with reference to Figure 2. The degradation function is stored in a memory 113 in one embodiment of the present invention. At the beginning and end of the sequence, there is insufficient data to define a context. Hence, the first k symbols and the last k' symbols are set to the corresponding symbols in the r sequence from the preliminary denoising system.

It should be noted that the received signals z and r do not need to be stored in a high-speed memory. At any given time, controller 111 during the first pass needs L symbols from z , and only one symbol from r . Hence, the received signal can be stored on a disk drive with the exception of a small buffer for storing the L symbols currently being utilized. Only the context memory 114 needs to be a high-speed memory.

The above examples assume a value for L has been determined. The present invention provides the greatest benefits in those cases in which the received sequence z has reference subsequences that are repeated a statistically significant number of times so that the counter values corresponding to any such subsequence lead to an accurate characterization of the behavior of the preliminary denoiser. If the number of observed occurrences of the

reference subsequences in the received sequence is small, the accuracy of the $N(J)$ counts discussed above might be low, and hence, the accuracy of the estimates $D(K)$ will likewise be low. If the accuracy of these counts is sufficiently low, the wrong decision with respect to correct output symbol will be made.

5

The number of occurrences of a reference subsequence depends to some degree on the length of the context. Consider the case in which a symbol z having a context of length $L-1$ is to be processed as described above. Further assume that the corresponding reference subsequence **azb**, appears Q times where $Q \gg 1$ and $Q/M \gg 1$, but the longer reference

10 subsequence **tazb** does not appear frequently for any value of t . Then a reference subsequence that is larger than L will have much fewer occurrences, and the statistical accuracy of the counts will be degraded relative to the case in which the smaller context was used. Hence, choosing too large a value for L can result in decision errors.

15

For any fixed L , the system can only exploit correlations among L samples or fewer in the input signal. The greater the extent of the input correlation that can be effectively exploited the better the performance. In contrast to the above considerations, this argues against making L too small.

20

From the above discussion, it is clear that there is an optimum value of L . This optimum can be determined empirically. If the length of the correlated sequences in the input signal does not change markedly over time, an optimum value for L can be determined experimentally by utilizing exemplary input signals and comparing the results of denoising for various values of L .

25

In principle, L can be determined for any particular output signal by denoising the signal using a number of different L values. In such a system, the value of L can be decreased from some upper bound until a value that provides satisfactory statistical accuracy is found. A reasonable starting value for L is given by $\lceil \log(n)/\log(M) \rceil$, where n is the number of

30 symbols in the z sequence and M is the number of symbols in the alphabet.

Refer now to Figure 3, which is a flow chart of the signal processing algorithm used in the second pass. After the received signal has been stored in memory 113, controller 111

sequentially examines the received symbols to determine if a symbol should be reset to another value. When controller 111 is at z_j , controller 111 reads the k symbols on the left of z_j and the k' symbols to the right of z_j to determine the largest reference subsequence $z(j)$ for z_j for which counts have been stored in memory 114 as shown at 151. Controller 111 extracts
 5 the counts associated with this reference subsequence from memory 114 as shown at 152 and determines if the stored counts have sufficient statistical accuracy to proceed as shown at 153. If the counts have sufficient accuracy, controller 111 reads the counts stored with $z(j)$ and estimates the signal degradation that would occur if z_j is replaced by each possible symbol value as shown at 155 utilizing Eq. (1) discussed above. The symbol is then set to the value
 10 that minimized the degradation as shown at 156.

If the statistical accuracy of the counts for this reference subsequence is too low, controller 111 looks for a smaller context as shown at 158. If such a context is present, the associated reference subsequence is chosen and the process repeated as shown at 160 and
 15 152. If no smaller context is available, z'_j is set to r_j , i.e., the value provided by the preliminary denoising system as shown at 159. The process continues by incrementing j as shown at 157 and repeating the process until all of the symbols that are to be processed have been processed. As noted above, the symbols on the ends of the sequence z' that are too close to an end to have a context are set to the values in the corresponding positions in the sequence
 20 r .

The above-described embodiments of the present invention have utilized a denoising apparatus that directly processes the received signal and has specific memories for use in storing the various parameters, contexts, and degradation functions. However, the present
 25 invention can be practiced on a general-purpose data processing system to which a copy of the received signal from the channel and a copy of the output of the preliminary denoising system have been transferred by loading an appropriate data processing program into that data processing system. Embodiments in which the preliminary denoising system operates on the same data processing system can also be practiced.

30

The above-described embodiments utilize separate memories for storing the degradation function, list of contexts, and the received signals. However, embodiments in which a single memory is used to store two or more of these quantities can also be

constructed without departing from the teachings of the present invention. Accordingly, it is to be understood that the separate memories discussed above can be part of a larger memory.

5 Various modifications to the present invention will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Accordingly, the present invention is to be limited solely by the scope of the following claims.